Entanglement-Assisted Capacity of Quantum Multiple Access Channels

Min-Hsiu Hsieh, Igor Devetak
Department of Electrical Engineering, University of Southern California, Los Angeles, California 90089, USA

Andreas Winter
Department of Mathematics, University of Bristol, Bristol BS8 1TW, U.K.

February 1, 2008

Abstract

We find a regularized formula for the entanglement-assisted (EA) capacity region for quantum multiple access channels (QMAC). We illustrate the capacity region calculation with the example of the collective phase-flip channel which admits a single-letter characterization. On the way we provide a first principles proof of the EA coding theorem based on a packing argument. We observe that the Holevo-Schumacher-Westmoreland theorem may be obtained from a modification of our EA protocol. We remark on the existence of a family hierarchy of protocols for multiparty scenarios with a single receiver, in analogy to the two-party case. In this way we relate several previous results regarding QMACs.

1 Introduction

Shannon's classical channel capacity theorem is one of the central results of information theory [18]. The capacity of a channel $(\mathcal{X}, Q(y|x), \mathcal{Y})$ is the the maximum rate at which classical information can be transmitted through a channel. It is given in terms of the mutual information between the channel input X and output Y

$$C = \max I(X;Y).$$

The maximization is performed over distributions $p(x)$ for the input random variables X, which together with the channel $p(y|x)$ determine the joint distribution for XY.

The multiple-access (MAC) channel $(\mathcal{X}_1 \times \mathcal{X}_2, p(y|x_1,x_2), \mathcal{Y})$ is a classical channel with two senders and one receiver. A general overview of MACs can be found in [4, 5]. The capacity problem now involves optimizing the information transmission rates R_1 and R_2 for each sender. The classical *capacity region* of a MAC was found independently by Ahlswede [1] and Liao [15]. It is given by the closure of the convex hull of all (R_1, R_2) satisfying

$$\begin{aligned} R_1 &\leq I(X_1;Y|X_2) \\ R_2 &\leq I(X_2;Y|X_1) \\ R_1 + R_2 &\leq I(X_1X_2;Y) \end{aligned} \tag{1}$$

for some product distribution $p(x_1)p(x_2)$ on $\mathcal{X}_1 \times \mathcal{X}_2$.

The *entanglement-assisted capacity* C_E of a *quantum channel* $\mathcal{N} : A' \to B$ is the maximum rate at which classical information can be transmitted through a channel if the sender Alice and

receiver Bob pre-share unlimited entanglement. Bennett, Shor, Smolin and Thapliyal (BSST) [2] (see also [13]) found a remarkably simple formula for the entanglement-assisted (EA) capacity in terms of the quantum mutual information between quantum systems A and B

$$C_E = \max I(A;B)_\theta. \tag{2}$$

The maximization is performed over states θ^{AB} of the form

$$\theta^{AB} = (I^A \otimes \mathcal{N})|\varphi\rangle\langle\varphi|^{AA'}.$$

Observe that this formula has the same form as Shannon's classical expression.

The main result of the present work regards *quantum* multiple access channels (QMAC). The QMAC was previously studied in [21, 23, 14]. We find an expression for the EA capacity region for the QMAC which has the same form as (1). As is common in quantum information theory [19], our formula is "regularized",which means that it is not efficiently computable. We exhibit a non-trivial example for which the capacity region can be efficiently computed.

The BSST proof of (2) employs a double blocking strategy based on the Holevo-Schumacher-Westmoreland (HSW) theorem [12, 17]. Our proof uses a first principles packing argument which perhaps sheds more light on the origin of the formula (2). The quantum mutual information $I(A;B)$ comes about from the "packing" of typical spaces of size $2^{nH(AB)}$ into the tensor product of two typical spaces of respective sizes $2^{nH(A)}$ and $2^{nH(B)}$. In addition, we recover the HSW theorem as a special case of the EA capacity theorem.

The paper is organized as follows. Section 2 contains the relevant background material. This includes notational conventions, definitions of the method of types, frequency typical sequences and subspaces, and useful lemmas. Section 3 contains statements and proofs of our main results. In section 4 we compute the capacity region of the collective phase-flip multiple access channel [23]. In section 5 we conclude by first rewriting our results in the resource inequality framework, from which we recover previously known coding theorems for the QMAC.

2 Background

If the state ρ is defined on the quantum system A we may denote it by ρ^A. We abuse notation and use A to denote the Hilbert space corresponding to the system A, as well as the set of bounded linear operators on this Hilbert space. We always use τ^A to denote the maximally mixed state $\tau^A = (\dim A)^{-1} I^A$ of a system A. We write the density operator of a pure state $|\psi\rangle$ as $\psi \equiv |\psi\rangle\langle\psi|$.

A quantum channel $\mathcal{N} : A' \to B$ is a cptp (completely positive trace preserving) map. It may be modelled by an isometry $U_\mathcal{N} : A' \to BE$ with a larger target space BE, followed by tracing out the "environment" system E. $U_\mathcal{N}$ is known as the Stinespring dilation of $\mathcal{N}$. We will often write $U_\mathcal{N}(\rho)$ for $U_\mathcal{N}\, \rho\, U_\mathcal{N}^\dagger$.

A quantum instrument [6] $\mathbf{D} = (\mathcal{D}_m)$, is an ordered set of cp (completely positive) maps $\mathcal{D}_m$,

$$\mathcal{D}_m : \rho \mapsto \sum_k A_{km} \rho A_{km}^\dagger.$$

The sum of the cp maps $\mathcal{D} = \sum_{m \in [\mu]} \mathcal{D}_m$ is trace preserving, $\sum_{km} A_{km}^\dagger A_{km} = I$. The instrument has one quantum input and two outputs, classical and quantum. The probability of classical outcome m and corresponding quantum output $\mathcal{D}_m(\rho)/(\mathrm{Tr}\,\mathcal{D}_m(\rho))$ is $\mathrm{Tr}\,\mathcal{D}_m(\rho)$. Ignoring the classical output reduces the instrument to the quantum map $\mathcal{D}$. Ignoring the quantum output reduces the instrument to the POVM (positive operator valued measure) (Λ_m) with $\Lambda_m = \sum_k A_{km}^\dagger A_{km}$.

The trace distance is defined as the trace norm of the difference between the two states

$$\|\sigma - \rho\|_1 = \mathrm{Tr}\,\sqrt{(\sigma-\rho)^2} = \max_{-I \le \Lambda \le I} \mathrm{Tr}\,[\Lambda(\sigma - \rho)].$$

The method of types is a powerful technique used in information theory. Denote by x^n a sequence $x_1 x_2 \dots x_n$, where each x_i belongs to the finite set $\mathcal{X}$. Denote by $|\mathcal{X}|$ the cardinality of $\mathcal{X}$. Denote by $N(x|x^n)$ the number of occurrences of the symbol x in the sequence x^n. The *type* t^{x^n} of a sequence x^n is a probability vector with elements $t_x^{x^n} = \frac{N(x_i|x^n)}{n}$. Denote the set of sequences of type t by

$$\mathcal{T}_t^n = \{x^n \in \mathcal{X}^n : t^{x^n} = t\}.$$

For the probability distribution p on the set $\mathcal{X}$ and $\delta > 0$, let $\tau_\delta = \{t : \forall x \in \mathcal{X},\ |t_x - p_x| \leq \delta\}$.Define the set of δ-typical sequences of length n as

$$\begin{aligned} \mathcal{T}_{p,\delta}^n &= \bigcup_{t \in \tau_\delta} \mathcal{T}_t^n \\ &= \{x^n : \forall x \in \mathcal{X},\ |t_x^{x^n} - p_x| \leq \delta\}. \end{aligned} \tag{3}$$

Define the probability distribution p^n on $\mathcal{X}^n$ to be the tensor power of p. The sequence x^n is drawn from p^n if and only if each letter x_i is drawn independently from p. Typical sequences enjoy many useful properties [4, 5]. Let $H(p) = -\sum_x p_x \log p_x$ be the Shannon entropy of p. For any $\epsilon, \delta > 0$, and all sufficiently large n for which

$$p^n(\mathcal{T}_{p,\delta}^n) \geq 1 - \epsilon \tag{4}$$

$$2^{-n[H(p)+c\delta]} \leq p^n(x^n) \leq 2^{-n[H(p)-c\delta]},\ \ \forall x^n \in \mathcal{T}_{p,\delta}^n \tag{5}$$

$$|\mathcal{T}_{p,\delta}^n| \leq 2^{n[H(p)+c\delta]}, \tag{6}$$

for some constant c. For $t \in \tau_\delta$ and for sufficiently large n, the cardinality D_t of $\mathcal{T}_t^n$ is bounded as [4]

$$D_t \geq 2^{n[H(p)-\eta(\delta)]} \tag{7}$$

and the function $\eta(\delta) \to 0$ as $\delta \to 0$.

The above concepts generalize to the quantum setting by virtue of the spectral theorem. Let $\rho = \sum_{x \in \mathcal{X}} p_x |x\rangle\langle x|$ be the spectral decomposition of a given density matrix ρ. In other words, $|x\rangle$ is the eigenstate of ρ corresponding to eigenvalue p_x. The von Neumann entropy of the density matrix ρ is

$$H(\rho) = -\mathrm{Tr}\, \rho \log \rho = H(p).$$

Define the type projector

$$\Pi_t^n = \sum_{x^n \in \mathcal{T}_t^n} |x^n\rangle\langle x^n|.$$

The density operator proportional to the type projector is $\tau_t = {D_t}^{-1} \Pi_t^n$. The typical subspace associated with the density matrix ρ is defined as

$$\Pi_{\rho,\delta}^n = \sum_{x^n \in \mathcal{T}_{p,\delta}^n} |x^n\rangle\langle x^n| = \sum_{t \in \tau_\delta} \Pi_t^n.$$

Properties analogous to (4) – (7) hold. For any $\epsilon, \delta > 0$, and all sufficiently large n for which

$$\mathrm{Tr}\, \rho^{\otimes n} \Pi_{\rho,\delta}^n \geq 1 - \epsilon \tag{8}$$

$$2^{-n[H(\rho)+c\delta]} \Pi_{\rho,\delta}^n \leq \Pi_{\rho,\delta}^n \rho^{\otimes n} \Pi_{\rho,\delta}^n \leq 2^{-n[H(\rho)-c\delta]} \Pi_{\rho,\delta}^n, \tag{9}$$

$$\operatorname{Tr}\Pi_{\rho,\delta}^{n} \leq 2^{n[H(\rho)+c\delta]}, \tag{10}$$

for some constant c. For $t \in \tau_\delta$ and for sufficiently large n, the support dimension of the type projector Π_t^n is bounded as

$$\operatorname{Tr}\Pi_t^n \geq 2^{n[H(\rho)-\eta(\delta)]}. \tag{11}$$

For a multipartite state ρ^{ABC} we write $H(A)_\rho = H(\rho^A)$, etc. We omit the subscript if the state is clear from the context. Define the quantum mutual information by

$$I(A;B) = H(A) + H(B) - H(AB)$$

and the quantum conditional mutual information by

$$I(A;C|B) = H(AB) + H(BC) - H(ABC) - H(B).$$

These are non-negative by strong subadditivity [16]. If $I(A;B) = 0$ then

$$I(A;C|B) = I(A;CB)$$

is easy to verify.

2.1 Useful Results

The set of generalized Pauli matrices $(U_m)_{m\in[d^2]}$, $[d^2] := \{1, \ldots, d^2\}$ is defined by $U_{l\cdot d+k} = \hat{Z}_d(l)\hat{X}_d(k)$ for $k, l = 0, 1, \cdots, d-1$ and

$$\begin{aligned} \hat{X}_d(k) &= \sum_s |s\rangle\langle s+k| = \hat{X}_d(1)^k, \\ \hat{Z}_d(l) &= \sum_s e^{i2\pi sl/d}|s\rangle\langle s| = \hat{Z}_d(1)^l. \end{aligned} \tag{12}$$

The $+$ sign denotes addition modulo d.

Define Φ^{AB} to be the maximally entangled state on a pair of d-dimensional systems A and B

$$|\Phi\rangle^{AB} = \frac{1}{\sqrt{d}}\sum_{i=0}^{d-1}|i\rangle^A|i\rangle^B. \tag{13}$$

In [2] it was shown that

$$\frac{1}{d^2}\sum_{m=1}^{d^2}(U_m \otimes I)\Phi^{AB}(U_m^\dagger \otimes I) = \tau^A \otimes \tau^B, \tag{14}$$

where $\tau = \frac{1}{d}I$. We will also use the property

$$(I \otimes U)|\Phi\rangle = (U^T \otimes I)|\Phi\rangle \tag{15}$$

for any operator U. T denotes transposition.

Next is a coherent version of the gentle operator lemma [20], Lemma 9. It states that a measurement which is likely to be successful in identifying a state tends not to significantly disturb the state.

Lemma 1 (Gentle coherent measurement) *Let* $(\rho_k^A)_{k\in[K]}$ *be a collection of density matrices and* $(\Lambda_k^A)_{k\in[K]}$ *a POVM such that*

$$\operatorname{Tr}\rho_k\Lambda_k \geq 1-\epsilon$$

for all k*. Let* $|\phi_k\rangle^{RA}$ *be a purification of* ρ_k^A*. Then there exists an isometric quantum operation* $\mathcal{D}: A \to AJ$ *such that*

$$\|(I^R\otimes\mathcal{D})(\phi_k^{RA}) - \phi_k^{RA}\otimes|k\rangle\langle k|^J\|_1 \leq \sqrt{8\epsilon}.$$

Proof Every POVM can be written as an isometry followed by projective measurement on a subsystem. In particular, there exists an isometry $\mathcal{D}: A\to AJ$ such that

$$(I^R\otimes\mathcal{D})|\phi\rangle^{RA} = \sum_j [(I^R\otimes\sqrt{\Lambda_j})|\phi\rangle^{RA}]|j\rangle^J.$$

Thus

$$\begin{aligned}
\langle k|\langle\phi_k|(I\otimes\mathcal{D})|\phi_k\rangle &= \langle\phi_k|(I\otimes\sqrt{\Lambda_k})|\phi_k\rangle \\
&\geq \langle\phi_k|(I\otimes\Lambda_k)|\phi_k\rangle \\
&= \operatorname{Tr}\rho_k\Lambda_k \\
&\geq 1-\epsilon.
\end{aligned} \tag{16}$$

The first inequality uses that $\Lambda_k \leq \sqrt{\Lambda_k}$ when $0\leq\Lambda_k\leq I$. The statement of the lemma follows from the fact that for pure states $|\zeta\rangle$ and $|\psi\rangle$,

$$\|\zeta-\psi\| = 2\sqrt{1-|\langle\zeta|\psi\rangle|^2}.$$

Lemma 2 (Packing) *We are given an ensemble* $\{\lambda_m, \sigma_m : m\in\mathcal{S}\}$ *with average density operator* $\sigma = \sum_{m\in\mathcal{S}}\lambda_m\sigma_m$*. Assume the existence of projectors* Π *and* $(\Pi_m)_{m\in\mathcal{S}}$ *with the following properties:*

$$\begin{aligned}
\operatorname{Tr}\sigma_m\Pi_m &\geq 1-\epsilon, && (17)\\
\operatorname{Tr}\sigma_m\Pi &\geq 1-\epsilon, && (18)\\
\operatorname{Tr}\Pi_m &\leq d, && (19)\\
\Pi\sigma\Pi &\leq D^{-1}\Pi, && (20)
\end{aligned}$$

for all $m\in\mathcal{S}$*. Let* $N = \lfloor\gamma D/d\rfloor$ *for some* $0<\gamma<1$*. Then there exists a map* $f:[N]\to\mathcal{S}$*, and a corresponding POVM* $(\Lambda_k)_{k\in[N]}$ *which reliably distinguishes between the states* $(\sigma_{f(k)})_{k\in[N]}$ *in the sense that*

$$\operatorname{Tr}\sigma_{f(k)}\Lambda_k \geq 1-4(\epsilon+\sqrt{8\epsilon})-16\gamma$$

for all $k\in[N]$*.*

The proof is in Appendix A.

Lemma 3 *If* $|\psi\rangle^{ABE}$ *is a pure state then*

$$H(B|E)_\psi = -H(B|A)_\psi$$

Proof Since $|\psi\rangle^{ABE}$ is pure, we have $H(A)_\psi = H(BE)_\psi$ and $H(E)_\psi = H(AB)_\psi$. Then

$$\begin{aligned}
H(B|E)_\psi &= H(BE)_\psi - H(E)_\psi \\
&= H(A)_\psi - H(AB)_\psi \\
&= -H(B|A)_\psi.
\end{aligned} \tag{21}$$

Lemma 4 *For any state σ^{ABE},*

$$I(A;B)_\sigma \leq H(B)_\sigma + H(B|E)_\sigma.$$

Proof Introduce a reference system R that purifies the state σ^{ABE}, then

$$\begin{aligned} I(A;B)_\sigma &= H(B)_\sigma - H(B|A)_\sigma \\ &= H(B)_\sigma + H(B|ER)_\sigma \\ &\leq H(B)_\sigma + H(B|E)_\sigma \end{aligned} \tag{22}$$

using lemma 3 and the fact that conditioning reduces entropy [16].

3 Main Result

3.1 Two party entanglement-assisted coding

Before attacking the multiuser problem we prove a version of the entanglement-assisted coding theorem. This theorem was first proved in [2] and subsequently in [13]. Both proofs invoke the Holevo-Schumacher-Westmoreland (HSW) theorem. The HSW theorem uses the method of conditionally typical subspaces. We give a direct proof based on the packing lemma which only uses typical subspaces. The proof perhaps sheds more light on why achievable rates take on the form of mutual information.

Alice and Bob are connected by a large number n uses of the quantum channel $\mathcal{N} : A' \to B$. Alice controls the channel input system A' and Bob has access to the channel output B. They also have entanglement in the form of n copies of some pure bipartite state $\varphi^{A'B'}$. Any such state is determined upto a local unitary transformation by the local density operator $\rho^{A'} = \mathrm{Tr}_{B'}\varphi^{A'B'}$. Alice and Bob use these resources to communicate, in analogy to superdense coding [3]. Based on her message Alice performs a quantum operation on her share of the entanglement. She then sends it through the quantum channel. Bob performs a decoding measurement on the channel output plus his share of the entanglement. They endeavour to maximize the communication rate.

We formalize the above information processing task. Define a $[n, R, \rho, \epsilon]$ entanglement-assisted code by

- a set of unitary encoding maps $(\mathcal{E}_k)_{k\in[2^{nR}]}$ acting on ${A'}^n = A'_1 \dots A'_n$ for Alice;
- Bob's decoding instrument $\mathbf{D} = (\mathcal{D}_k)_{k\in[2^{nR}]}$ acting on $B'^n B^n$.

such that for all $k \in [2^{nR}]$

i) $\mathrm{Tr}\,\{[\mathcal{D}_k \circ ((\mathcal{N}^{\otimes n} \circ \mathcal{E}_k) \otimes I)](\varphi^{\otimes n})\} \geq 1 - \epsilon$;

ii) the code density operator satisfies $\mathcal{E}_k(\rho^{\otimes n}) = \rho^{\otimes n}$;

iii) $\left\|[(\mathcal{D} \otimes I^{E^n}) \circ ((U_{\mathcal{N}}{}^{\otimes n} \circ \mathcal{E}_k) \otimes I) - (U_{\mathcal{N}}{}^{\otimes n} \otimes I)](\varphi^{\otimes n})\right\|_1 \leq \epsilon$.

Condition i) means that Bob correctly decodes Alice's message with high probability. This condition suffices for two-party entanglement-assisted communication. The other two properties will prove important for the multiparty generalization. Condition ii) means that Alice always inputs a tensor power state into the channel. Condition iii) says that the encoding and decoding operations in effect cancel each other out.

Theorem 1 *Define $\theta^{AB} = (I \otimes \mathcal{N})\varphi^{AA'}$ and $R = I(A;B)_\theta$. For every $\epsilon, \delta > 0$ and n sufficiently large, there exists an $[n, R - \delta, \rho, \epsilon]$ entanglement-assisted code.*

Proof of Theorem 1 Let $t(1), \ldots, t(a)$ be an ordering of the distinct types t^{x^n}. Define $\tau_\alpha^n = 1/d_\alpha \Pi_{t(\alpha)}^n$, $d_\alpha = \operatorname{Tr}\Pi_{t(\alpha)}^n$ and

$$|\Phi_\alpha\rangle^{A'^n B'^n} = \frac{1}{\sqrt{d_\alpha}} \sum_{x^n \in \mathcal{T}_{t(\alpha)}^n} |x^n\rangle^{A'^n} |x^n\rangle^{B'^n}.$$

In the beginning Alice and Bob share the entangled state

$$\begin{aligned} |\Psi\rangle^{A'^n B'^n} &= |\varphi\rangle^{\otimes n} \\ &= \sum_\alpha \sqrt{p_\alpha} |\Phi_\alpha\rangle, \end{aligned} \tag{23}$$

where $p_\alpha = \sum_{x^n \in \mathcal{T}_{t(\alpha)}^n} p^n(x^n)$. The type projectors $\Pi_{t(\alpha)}^n$ induce a decomposition of the Hilbert space $\mathcal{H}^{\otimes n}$ of A'^n and B'^n into a direct sum

$$\mathcal{H}^{\otimes n} = \bigoplus_{\alpha=1}^{a} \mathcal{H}_{t(\alpha)}.$$

Let $\mathcal{G} = \{(g_1, g_2, \cdots, g_a) : g_\alpha \in \{1, \ldots, d_\alpha^2\}\}$, $\mathcal{B} = \{(b_1, b_2, \cdots, b_a) : b_\alpha \in \{0,1\}\}$, and $\mathcal{S} = \mathcal{G} \times \mathcal{B}$. Every element $s^a \in \mathcal{S}$ is uniquely determined by $g^a \in \mathcal{G}$ and $b^a \in \mathcal{B}$. Given an element $s^a \in \mathcal{S}$, define a unitary operation U_{s^a} to be

$$U_{s^a} \equiv U_{g^a, b^a} = \bigoplus_{\alpha=1}^{a} (-1)^{b_\alpha} U_{g_\alpha} \tag{24}$$

where $\{U_{g_\alpha}\}$ are the d_α^2 generalized Pauli operators (12) defined on $\mathcal{H}_{t(\alpha)}$. Define

$$\begin{aligned} \sigma_{s^a}^{B^n B'^n} &:= (\mathcal{N}^{\otimes n} \otimes I) \left[(U_{s^a} \otimes I) \Psi^{A'^n B'^n} (U_{s^a}^\dagger \otimes I) \right] \\ &= (I \otimes U_{s^a}^T) \theta^{\otimes n} (I \otimes U_{s^a}^*). \end{aligned} \tag{25}$$

The last equality follows from (15). Let σ to be the average of σ_{s^a} over $\mathcal{S}$, then

$$\begin{aligned} \sigma &= \frac{1}{|\mathcal{S}|} \sum_{s^a \in \mathcal{S}} \sigma_{s^a} \\ &= \frac{1}{|\mathcal{B}||\mathcal{G}|} \sum_{g^a \in \mathcal{G}} \sum_{b^a \in \mathcal{B}} \sum_{\alpha, \alpha'} \sqrt{p_\alpha p_{\alpha'}} (\mathcal{N}^{\otimes n} \otimes I) \left[(U_{g^a, b^a} \otimes I) |\Phi_\alpha\rangle\langle\Phi_{\alpha'}| (U_{g^a, b^a}^\dagger \otimes I) \right]. \\ &= \sum_\alpha p_\alpha \Big(\mathcal{N}^{\otimes n}(\tau_\alpha^n) \otimes \tau_\alpha^n \Big). \end{aligned} \tag{26}$$

The last equality comes from (27) and (28) below. When $\alpha = \alpha'$,

$$\begin{aligned} &\frac{1}{|\mathcal{B}||\mathcal{G}|} \sum_{g^a \in \mathcal{G}} \sum_{b^a \in \mathcal{B}} p_\alpha (\mathcal{N}^{\otimes n} \otimes I) \left[(U_{g^a, b^a} \otimes I) \Phi_\alpha (U_{g^a, b^a}^\dagger \otimes I) \right] \\ &= (\mathcal{N}^{\otimes n} \otimes I) \frac{1}{|\mathcal{G}|} \sum_{g_1} \cdots \sum_{g_a} p_\alpha (U_{g_\alpha} \otimes I) \Phi_\alpha (U_{g_\alpha}^\dagger \otimes I) \\ &= (\mathcal{N}^{\otimes n} \otimes I) p_\alpha (\tau_\alpha^n \otimes \tau_\alpha^n). \end{aligned} \tag{27}$$

The last equality follows from (14). When $\alpha \neq \alpha'$,

$$\begin{aligned}
&\frac{1}{|\mathcal{B}||\mathcal{G}|} \sum_{g^a \in \mathcal{G}} \sum_{b^a \in \mathcal{B}} \sqrt{p_\alpha p_{\alpha'}} (\mathcal{N}^{\otimes n} \otimes I) \left[(U_{g^a, b^a} \otimes I) |\Phi_\alpha\rangle\langle\Phi_{\alpha'}| (U^\dagger_{g^a, b^a} \otimes I) \right] \\
&= \frac{1}{d_\alpha^2 d_{\alpha'}^2} \sqrt{p_\alpha p_{\alpha'}} \sum_{b_\alpha b_{\alpha'}} \frac{(-1)^{b_\alpha + b_{\alpha'}}}{4} \left\{ \sum_{g_\alpha g_{\alpha'}} (\mathcal{N}^{\otimes n} \otimes I) \left[(U_{g_\alpha} \otimes I) |\Phi_\alpha\rangle\langle\Phi_{\alpha'}| (U^\dagger_{g_{\alpha'}} \otimes I) \right] \right\} \\
&= 0.
\end{aligned} \tag{28}$$

Define the projectors on $B'^n B^n$

$$\begin{aligned}
\Pi_{s^a} &= (I \otimes U^T_{s^a})\, \Pi^n_{\theta,\delta}\, (I \otimes U^*_{s^a}), \\
\Pi &= \Pi^n_{\mathcal{N}(\rho),\delta} \otimes \Pi^n_{\rho,\delta}.
\end{aligned} \tag{29}$$

The following properties are proved in Appendix B. For all $\epsilon > 0, \delta > 0$ and all sufficiently large n,

$$\operatorname{Tr} \sigma_{s^a} \Pi \geq 1 - \epsilon \tag{30}$$

$$\operatorname{Tr} \sigma_{s^a} \Pi_{s^a} \geq 1 - \epsilon \tag{31}$$

$$\operatorname{Tr} \Pi_{s^a} \leq 2^{n[H(AB)_\theta + c\delta]} \tag{32}$$

$$\Pi \sigma \Pi \leq 2^{n[H(A)_\theta + H(B)_\theta + c\delta]} \Pi. \tag{33}$$

Let $\lambda_{s^a} = \frac{1}{|\mathcal{S}|}$ and $R = I(A;B)_\theta - (2c+1)\delta$. We now apply the packing lemma to the ensemble $\{\lambda_{s^a}, \sigma_{s^a} : s^a \in \mathcal{S}\}$ and projectors Π and Π_{s^a}. Thus there exist a map $f : [2^{nR}] \to \mathcal{S}$ and a POVM $\{\Lambda_k\}_{k \in [2^{nR}]}$ such that

$$\operatorname{Tr} \sigma_{f(k)} \Lambda_k \geq 1 - \epsilon', \tag{34}$$

with

$$\epsilon' = 4(\epsilon + \sqrt{8\epsilon}) + 16 \times 2^{-n\delta}.$$

Define the encoding operation by $\mathcal{E}_k = U_{f(k)}$.

Including the environment system, the state of $B^n B'^n E^n$ after the application of the channel $U_\mathcal{N}$ is

$$\begin{aligned}
|\Upsilon_k\rangle^{B^n B'^n E^n} &= ({U_\mathcal{N}}^{\otimes n} \otimes I)(U_{f(k)} \otimes I)|\Psi\rangle^{A'^n B'^n} \\
&= ({U_\mathcal{N}}^{\otimes n} \otimes U^T_{f(k)})|\Psi\rangle^{A'^n B'^n}
\end{aligned} \tag{35}$$

$|\Upsilon_k\rangle$ is a purification of $\sigma_{f(k)}$. By Lemma 1, there exists an isometry $\mathcal{D}' : B^n B'^n \to B^n B'^n J$ such that

$$\|(I \otimes \mathcal{D}')(\Upsilon_k) - \Upsilon_k \otimes |k\rangle\langle k|^J\|_1 \leq \sqrt{8\epsilon'}.$$

Bob performs the controlled unitary

$$W^{JB'^n} = \sum_k |k\rangle\langle k|^J \otimes (U^*_{f(k)})^{B'^n}.$$

Defining $\mathcal{D}'' = (W \otimes I^{B^n}) \circ \mathcal{D}'$, this implies

$$\|(I \otimes \mathcal{D}'')(\Upsilon_k) - [(U_\mathcal{N}^{\otimes n} \otimes I)(\varphi^{\otimes n})] \otimes |k\rangle\langle k|\|_1 \leq \sqrt{8\epsilon'}. \tag{36}$$

The instrument $(\mathcal{D}_k)$ is defined by $\mathcal{D}''$ followed by a von Neumann measurement of the system J. Equation (36) expresses the fact that the classical communication being performed is almost

decoupled from all the quantum systems involved in the protocol, including ancillas and the inaccessible environmnent. We remark that this guarnatees the ability to "coherify" the protocol in the sense of [7].

Condition i) in the form

$$\mathrm{Tr}\left\{[\mathcal{D}_k \circ ((\mathcal{N}^{\otimes n} \circ \mathcal{E}_k) \otimes I)](\varphi^{\otimes n})\right\} \geq 1 - \epsilon'$$

is immediate from (34). Condition ii) follows from the construction (24). Condition iii) in the form

$$\left\| [(\mathcal{D} \otimes I^{E^n}) \circ ((U_{\mathcal{N}}^{\otimes n} \circ \mathcal{E}_k) \otimes I) - (U_{\mathcal{N}}^{\otimes n} \otimes I)](\varphi^{\otimes n}) \right\|_1 \leq \sqrt{8\epsilon'}$$

follows from (36). ■

3.2 Remark on the HSW theorem

Suppose that Alice and Bob are connected by a special *cq channel* of the form

$$\mathcal{N} = \mathcal{N}' \circ \Delta,$$

where Δ is the dephasing channel

$$\rho \rightarrow \sum_x |x\rangle\langle x|\rho|x\rangle\langle x|.$$

A $\{c \rightarrow q\}$ channel is equivalent to one with classical inputs and quantum outputs. The HSW coding theorem states that rates $R = I(A;B)_\theta$, $\theta^{AB} = (I \otimes \mathcal{N})\varphi^{AA'}$ are achievable even without entanglement assistance. We show that this fact follows from our construction in two steps.

The first step is to replace the entanglement used by classical common randomness. Observe that the encoding operations U_{s^a} all satisfy

$$\Delta^{\otimes n} \circ U_{s^a} = \Delta^{\otimes n} \circ U_{s^a} \circ \Delta^{\otimes n}.$$

This follows from the corresponding property of the generalized Pauli operators (12). Hence for cq channels $\mathcal{N}$

$$\begin{aligned} \sigma_{f(k)} &= [(\mathcal{N}^{\otimes n} \circ \mathcal{E}_k) \otimes I](\varphi^{\otimes n}) \\ &= [(\mathcal{N}^{\otimes n} \circ \mathcal{E}_k \circ \Delta^{\otimes n}) \otimes I](\varphi^{\otimes n}) \\ &= [(\mathcal{N}^{\otimes n} \circ \mathcal{E}_k) \otimes I](\overline{\varphi}^{\otimes n}), \end{aligned} \tag{37}$$

where

$$\overline{\varphi} = (\Delta \otimes I)\varphi = \sum_x p_x |x\rangle\langle x| \otimes |x\rangle\langle x|$$

is the dephased version of φ. The state $\overline{\varphi}^{\otimes n}$ can be constructed from classical common randomness like that used in Shannon's original coding theorem.

The second step is showing that common randomness is not needed. The argument parallels the derandomization step from the proof of the packing lemma (Appendix A). We have thus recovered the HSW coding theorem.

The benefit of the above proof is its close analogy to Shannon's joint typicality decoding. We only made use of typical subspaces and not conditionally typical subspaces.

3.3 Multiple Access Channel

We turn to the communication scenario with two senders, Alice and Bob, and one receiver, Charlie. They are connected by a large number n of uses of the *multiple access* quantum channel $\mathcal{M} : A'B' \rightarrow C$. Alice and Bob controls the channel input systems A' and B', respectively. Charlie has access to the channel output C. Each sender also shares unlimited entanglement with the receiver, in the form of arbitrary pure states $|\Gamma_1\rangle^{AC_A}$ and $|\Gamma_2\rangle^{BC_B}$. The system A is held by Alice, B by Bob, and $C_A C_B$ by Charlie. Based on her message Alice performs a quantum operation on her share of the entanglement, and likewise for Bob. These are then sent through the quantum channel. Charlie performs a decoding measurement on the channel output plus his share of the entanglement. Now both Alice's and Bob's communication rates need to be optimized.

We formalize the above information processing task.

Define an (n, R_1, R_2, ϵ) entanglement-assisted code by

- two sets of encoding cptp maps: $(\mathcal{E}_k^1)_{k\in[2^{nR_1}]}$ taking A to A'^n for Alice, and $(\mathcal{E}_l^2)_{l\in[2^{nR_2}]}$ taking B to B'^n for Bob ;
- Charlie's decoding POVM $(\Lambda_{k,l})_{k\in[2^{nR_1}],l\in[2^{nR_2}]}$ on $C_A C_B C$,

such that

$$\mathrm{Tr}\,\{\Lambda_{k,l}[((\mathcal{M}^{\otimes n} \circ (\mathcal{E}_k^1 \otimes \mathcal{E}_l^2)) \otimes I^{C_A C_B})(\Gamma_1^{AC_A} \otimes \Gamma_2^{BC_B})]\} \geq 1 - \epsilon. \tag{38}$$

We say that (R_1, R_2) is an *achievable rate pair* if for all $\epsilon > 0, \delta > 0$ and sufficiently large n there exists an $(n, R_1 - \delta, R_2 - \delta, \epsilon)$ entanglement-assisted code. The entanglement-assisted *capacity region* $C_E(\mathcal{M})$ is defined to be the closure of the set of achievable rate pairs.

Theorem 2 *Consider a quantum multiple access channel $\mathcal{M} : A'B' \rightarrow C$. For some states $\rho_1^{A'}$ and $\rho_2^{B'}$ define*

$$\theta^{ABC} = (I^{AB} \otimes \mathcal{M})(\varphi_1^{AA'} \otimes \varphi_2^{BB'}), \tag{39}$$

where $|\varphi_1\rangle^{AA'}$ and $|\varphi_2\rangle^{BB'}$ are purifications of $\rho_1^{A'}$ and $\rho_2^{B'}$ respectively. Define the two-dimensional region $C_E(\mathcal{M}, \rho_1, \rho_2)$, shown in Fig.1, by the set of pairs of nonnegative rates (R_1, R_2) satisfying

$$\begin{aligned} R_1 &\leq I(A; C|B)_\theta \\ R_2 &\leq I(B; C|A)_\theta \\ R_1 + R_2 &\leq I(AB; C)_\theta. \end{aligned} \tag{40}$$

Define $\widetilde{C}_E(\mathcal{M})$ as the convex hull over states ρ_1, ρ_2 of $C_E(\mathcal{M}, \rho_1, \rho_2)$. Then the entanglement-assisted capacity region $C_E(\mathcal{M})$ is given by the regularized expression

$$C_E(\mathcal{M}) = \overline{\bigcup_{n=1}^{\infty} \widetilde{C}_E(\mathcal{M}^{\otimes n})}.$$

There is an additional achievable *single-letter upper bound on the sum rate*

$$R_1 + R_2 \leq \max_{\rho_1,\rho_2} I(AB; C)_\theta \tag{41}$$

Proof of Theorem 2 (direct coding theorem) Let the entanglement be given in a tensor power form, as in Theorem 1. Define a $[n, R_1, R_2, \rho_1, \rho_2, \epsilon]$ entanglement-assisted code as a special case of an (n, R_1, R_2, ϵ) code: specify $\Gamma_1 = \phi_1^{\otimes n}$ and $\Gamma_2 = \phi_2^{\otimes n}$, and identify $A := A'^n$ and $B := B'^n$.

To show the achievability of every rate pair (R_1, R_2) in the convex hull of the $C_E(\mathcal{M}, \rho_1, \rho_2)$, by time-sharing it suffices to show this for elements of a particular $C_E(\mathcal{M}, \rho_1, \rho_2)$. The latter follows from the achievability of the corner points in the region. Once we show that, the non-corner points

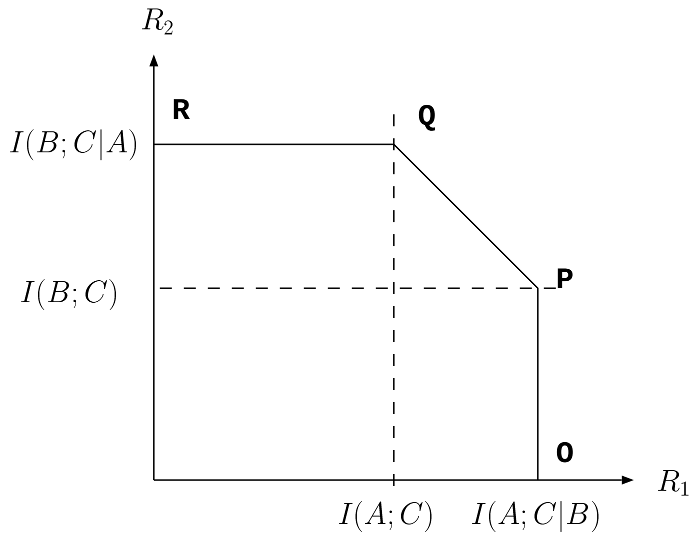

Figure 1: Capacity region of multiple access channel for fixed input states ρ_1 and ρ_2

can be achieved by time-sharing (see, e.g. [9]). Consider the corner point Q. For all $\epsilon > 0, \delta > 0$ and n sufficiently large, we show below that there exists a $[n, I(A;C)_\theta - \delta, I(B;C|A)_\theta - \delta, \rho_1, \rho_2, \epsilon]$ entanglement-assisted code $(\mathcal{E}^1, \mathcal{E}^2, \mathcal{D})$.

The point Q corresponds to the maximum rate that at which Alice can send as long as Bob sends at his maximum rate. This is the rate that is achieved when Bob's input is considered as noise for the channel from Alice to Charlie. From the two party direct coding theorem, Alice can send at a rate $I(A;C)$ and Charlie can decode the message with arbitrarily low probability. Charlie then knows which encoding operation Alice used and can subtract its effect from the channel. Therefore, Bob can achieve the rate $I(B;C|A)$. This outlines the proof of the achievability of point Q.

Define the channel $\mathcal{N}_1 : A' \to C$ by

$$\mathcal{N}_1 : \omega \mapsto \mathcal{M}(\omega \otimes \rho_2).$$

$\mathcal{N}_1^{\otimes n}$ is the effective channel from Alice to Charlie when Bob's input to $\mathcal{M}^{\otimes n}$ is $\rho_2^{\otimes n}$. Define $\hat{\mathcal{N}}_1 : A' \to C_B C$ by

$$\hat{\mathcal{N}}_1 : \omega \mapsto (I \otimes \mathcal{M})(\omega \otimes \varphi_2).$$

Observe that $\hat{\mathcal{N}}_1$ is an extension of $\mathcal{N}_1$. Hence it is a restriction of $U_{\mathcal{N}_1}$.

Define the channel $\mathcal{N}_2 : B' \to C_A C$ by

$$\mathcal{N}_2 : \omega \mapsto (I \otimes \mathcal{M})(\varphi_1 \otimes \omega).$$

$\mathcal{N}_2^{\otimes n}$ is effective the channel from Bob to Charlie if Alice simply inputs the A' part of the entangled state/purification $(\varphi_1^{A'C_A})^{\otimes n}$ without encoding.

Fix $\epsilon > 0, \delta > 0$. Define $R_1 = I(A;C)_\theta - \delta$ and $R_2 = I(B;C|A)_\theta - \delta$, with θ defined in (39). By Theorem 1, for sufficiently large n there exists an $[n, R_1, \rho_1, \epsilon]$ entanglement-assisted code $(\mathcal{E}^1, \mathcal{D}^1)$ for $\mathcal{N}_1$ and an $[n, R_2, \rho_2, \epsilon]$ entanglement-assisted code $(\mathcal{E}^2, \mathcal{D}^2)$ for $\mathcal{N}_2$ such that for all $k \in [2^{nR_1}], l \in [2^{nR_2}]$,

i) $\mathrm{Tr}\left\{[\mathcal{D}_k^1 \circ ((\mathcal{N}_1^{\otimes n} \circ \mathcal{E}_k^1) \otimes I^{C_A})](\varphi_1^{\otimes n})\right\} \geq 1 - \epsilon$;

ii) $\left\| [(\mathcal{D}^1 \otimes I) \circ ((\hat{\mathcal{N}}_1^{\otimes n} \circ \mathcal{E}_k^1) \otimes I^{C_A}) - (\hat{\mathcal{N}}_1^{\otimes n} \otimes I^{C_A})](\varphi_1^{\otimes n}) \right\|_1 \leq \epsilon$;

iii) $\mathrm{Tr}\,\{[\mathcal{D}_l^2 \circ ((\mathcal{N}_2^{\otimes n} \circ \mathcal{E}_l^2) \otimes I^{C_B})](\varphi_2^{\otimes n})\} \geq 1 - \epsilon$;

iv) the code density operator satisfies $\mathcal{E}_l^2(\rho_2^{\otimes n}) = \rho_2^{\otimes n}$.

We now define our code for the multiple access channel $\mathcal{M}$. Alice and Bob encode according to $(\mathcal{E}_k^1)$ and $(\mathcal{E}_l^2)$, respectively. Define the instrument $(\mathcal{D}_{k,l})$ on CC_AC_B by

$$\mathcal{D}_{k,l} = \mathcal{D}_l^2 \circ (\mathcal{D}_k^1 \otimes I^{C_B}).$$

Then Charlie's decoding POVM $(\Lambda_{k,l})$ is the restriction of $(\mathcal{D}_{k,l})$. Examining the success probability of decoding Alice's message k:

$$\begin{aligned}
&\mathrm{Tr}\,\{(\mathcal{D}_k^1 \otimes I^{C_B}) \circ ((\mathcal{M}^{\otimes n} \circ (\mathcal{E}_k^1 \otimes \mathcal{E}_l^2)) \otimes I^{C_AC_B})(\varphi_1^{\otimes n} \otimes \varphi_2^{\otimes n})\} \\
&= \mathrm{Tr}\,\{\mathcal{D}_k^1 \circ ((\mathcal{M}^{\otimes n} \circ (\mathcal{E}_k^1 \otimes \mathcal{E}_l^2)) \otimes I^{C_A})(\varphi_1^{\otimes n} \otimes \rho_2^{\otimes n})\} \\
&= \mathrm{Tr}\,\{\mathcal{D}_k^1 \circ ((\mathcal{M}^{\otimes n} \circ (\mathcal{E}_k^1 \otimes I^{B'^n})) \otimes I^{C_A})(\varphi_1^{\otimes n} \otimes \rho_2^{\otimes n})\} \\
&= \mathrm{Tr}\,\{[\mathcal{D}_k^1 \circ ((\mathcal{N}_1^{\otimes n} \circ \mathcal{E}_k^1) \otimes I)](\varphi_1^{\otimes n})\} \\
&\geq 1 - \epsilon.
\end{aligned} \tag{42}$$

The second equality follows from iv) and the third from i).

Next Charlie decodes Bob's message. Rewrite ii) in terms of $\mathcal{M}$:

$$\left\|[(\mathcal{D}^1 \otimes I^{C_B}) \circ ((\mathcal{M}^{\otimes n} \circ (\mathcal{E}_k^1 \otimes I^{B'^n})) \otimes I^{C_AC_B}) - (\mathcal{M}^{\otimes n} \otimes I^{C_AC_B})](\varphi_1^{\otimes n} \otimes \varphi_2^{\otimes n})\right\|_1 \leq \epsilon;$$

Since $\mathcal{E}_l^2$ is unitary and satisfies iv),

$$\left\|[(\mathcal{D}^1 \otimes I^{C_B}) \circ ((\mathcal{M}^{\otimes n} \circ (\mathcal{E}_k^1 \otimes \mathcal{E}_l^2)) \otimes I^{C_AC_B}) - ((\mathcal{M}^{\otimes n} \circ (I^{A'^n} \otimes \mathcal{E}_l^2)) \otimes I^{C_AC_B})](\varphi_1^{\otimes n} \otimes \varphi_2^{\otimes n})\right\|_1 \leq \epsilon;$$

Rewrite iii) in terms of $\mathcal{M}$:

$$\mathrm{Tr}\,\{[\mathcal{D}_l^2 \circ ((\mathcal{M}^{\otimes n} \circ (I^{A'^n} \otimes \mathcal{E}_l^2)) \otimes I^{C_AC_B})](\varphi_1^{\otimes n} \otimes \varphi_2^{\otimes n})\} \geq 1 - \epsilon;$$

Define

$$\Omega^{CC_AC_B} = (\mathcal{D}^1 \otimes I^{C_B}) \circ ((\mathcal{M}^{\otimes n} \circ (\mathcal{E}_k^1 \otimes \mathcal{E}_l^2)) \otimes I^{C_AC_B})(\varphi_1^{\otimes n} \otimes \varphi_2^{\otimes n})$$

Hence

$$\mathrm{Tr}\,[\mathcal{D}_l^2\, \Omega^{CC_AC_B}] \geq 1 - 2\epsilon.$$

Now (38) follows. This concludes the achievability of point Q.

Corner point P can be shown in the same manner. Corner point R corresponds to the maximum rate achievable from Bob to Charlie when Alice is not sending any information. The proof is obvious since we can assume that Alice is throwing the same state into the channel all the time. The corner point O follows from the same reasoning. This concludes the proof of direct coding theorem.

Remark. The entanglement assistance may be phrased in terms of tensor powers of ebit states $|\Phi_+\rangle = \frac{1}{\sqrt{2}}(|0\rangle|0\rangle + |1\rangle|1\rangle)$ instead of the arbitrary $|\Gamma_1\rangle$ and $|\Gamma_2\rangle$. The protocol achieving the corner points of the region $C_E(\mathcal{M}, \rho_1, \rho_2)$ uses $|\Gamma_1\rangle = |\phi_1\rangle^{\otimes n}$ and $|\Gamma_2\rangle = |\phi_2\rangle^{\otimes n}$. By entanglement dilution [10], $|\Gamma_1\rangle$ may be asymptotically obtained from an ebit rate of $E_1 = H(A)_\theta$ shared between Alice and Charlie. Likewise $|\Gamma_2\rangle$ may be asymptotically obtained from an ebit rate of $E_2 = H(B)_\theta$ shared between Bob and Charlie.

Proof of Theorem 2 (converse) Start with some (n, R_1, R_2, ϵ) entanglement-assisted code (see Fig. 2). Assume Alice's message k and Bob's message l are picked according to the uniform distributions on $[2^{nR_1}]$ and $[2^{nR_2}]$, respectively. These correspond to random variables K and L.

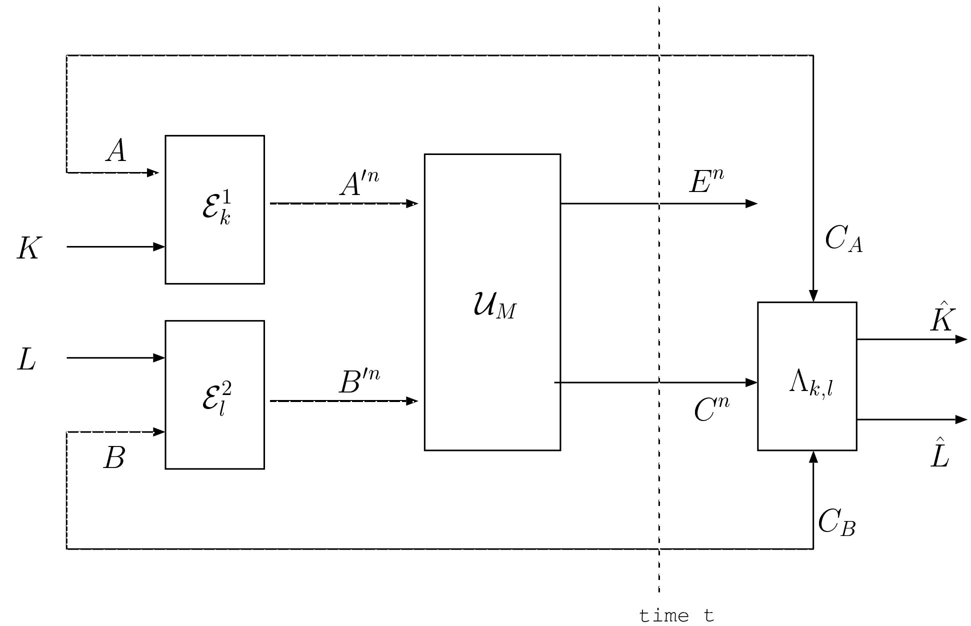

Figure 2: A general protocol for multiple access entanglement-assisted classical communication

Alice performs the encoding operation $\mathcal{E}_k^1$ on the A part of $|\Gamma_1\rangle^{AC_A}$ conditioned on $K = k$. Bob performs the encoding operation $\mathcal{E}_l^2$ on the B part of $|\Gamma_2\rangle^{BC_B}$ conditioned on $L = l$. The output of $\mathcal{E}_k^1 \otimes \mathcal{E}_l^2$ is sent through the multiple access channel $\mathcal{M}^{\otimes n}$ just after time t_0 . The channel output C^n is acquired by Charlie at time t. Charlie performs a POVM on the channel output and his part of the entanglement $C_A C_B$. The measurement outcome is a random variable $W = (\hat{K}, \hat{L})$. By the condition (38),

$$\Pr\{K \neq \hat{K} \,\text{and}\, L \neq \hat{L}\} \leq \epsilon. \tag{43}$$

The protocol ends at time t_f. We first obtain an upper bound on the sum rate $R_1 + R_2$. At this time

$$n(R_1 + R_2) = H(KL) \leq I(KL; \hat{K}\hat{L}) + n\eta(n, \epsilon), \tag{44}$$

where the function $\eta(n, \epsilon)$ tends to 0 as ϵ tends to 0 and n tends to infinity. The inequality is standard in classical information theory [4]. It is obtained by applying Fano's inequality [4] to (43). Denote the state of the system at time t by

$$\omega^{KLC_AC_BC^nE^n} = (I^{KLC_AC_B} \otimes U_{\mathcal{M}}^{\otimes n})(\xi_1 \otimes \xi_2),$$

$$\xi_1^{A'^nKC_A} = 2^{-nR_1} \sum_k |k\rangle\langle k|^K \otimes (\mathcal{E}_k^1 \otimes I^{C_A})(\Gamma_1^{AC_A}),$$

$$\xi_2^{B'^nLC_B} = 2^{-nR_2} \sum_l |l\rangle\langle l|^L \otimes (\mathcal{E}_l^2 \otimes I^{C_B})(\Gamma_2^{BC_B}).$$

Denote by A^n the system which purifies the restriction of the A'^n parts of the state ξ_1 at time t_0. Then A^n contains K and C_A as subsystems. Define B^n in a similar fashion.

The Holevo bound reads

$$I(KL; \hat{K}\hat{L}) \leq I(KL; C_AC_BC^n)_\omega. \tag{45}$$

The entropic quantities below refer to the state ω.

$$\begin{aligned}
&I(KL;C_AC_BC^n)\\
&=I(C^n;C_AC_BKL)-I(C_AC_B;C^n)+I(KL;C_AC_B)\\
&\le I(C^n;C_AC_BKL)\\
&\le H(C^n)+H(C^n|E^n)\\
&=H(C^n)-H(C^n|A^nB^n)\\
&=I(C^n;A^nB^n).
\end{aligned} \tag{46}$$

The first inequality follows from $I(KL;C_AC_B)=0$ and $I(C_AC_B;C^n)\ge 0$. The second inequality holds because of lemma 4. The second equality is from lemma 3.

Putting everything together gives

$$R_1+R_2\le \eta(n,\epsilon)+\frac{1}{n}I(C^n;A^nB^n). \tag{47}$$

Observe that

$$\begin{aligned}
&\frac{1}{n}H(C^n)+H(C^n|E^n)\\
&\le \frac{1}{n}\sum_i [H(C_i)+H(C_i|E_i)]\\
&\le \max_{\rho_1,\rho_2}[H(C)_\theta+H(C|E)_\theta]\\
&=\max_{\rho_1,\rho_2}[H(C)_\theta-H(C|AB)_\theta]\\
&=\max_{\rho_1,\rho_2} I(AB;C)_\theta.
\end{aligned} \tag{48}$$

The state θ is defined in (39).

An upper bound on Alice's rate R_1 is obtained in a similar fashion. Equations

$$nR_1=H(K)\le I(K;\hat{K})+n\eta(n,\epsilon), \tag{49}$$

and

$$I(K;\hat{K})\le I(K;C_AC_BC^n)_\omega \tag{50}$$

are obtained as above. With respect to ω:

$$\begin{aligned}
&I(K;C_AC_BC^n)\\
&=I(C_BC^n;C_AK)-I(C_A;C_BC^n)+I(K;C_A)\\
&\le I(C_BC^n;C_AK)\\
&\le I(B^nC^n;C_AK)\\
&\le H(B^nC^n)+H(B^nC^n|E^n)\\
&=H(B^nC^n)-H(B^nC^n|A^n)\\
&=I(A^n;B^nC^n)\\
&=I(A^n;C^n|B^n).
\end{aligned} \tag{51}$$

Hence

$$R_1\le \eta(n,\epsilon)+\frac{1}{n}I(A^n;C^n|B^n). \tag{52}$$

By the same argument

$$R_2\le \eta(n,\epsilon)+\frac{1}{n}I(B^n;C^n|A^n). \tag{53}$$

This concludes the proof of the converse.

We leave it as an open problem to single-letterize the above capacity region. We do not know if the regularization in our main theorem is actually necessary. Indictions that it might not be are the successful single-letterization of the two-user entanglement-assisted capacity in [2] which we have used to obtain the single-letter bound on the rate-sum above, and the fact that the regularization is not necessary in the classical case.

4 The collective phase-flip channel example

Consider the case that $|A'| = |B'| = d$. The collective phase-flip channel [23] $\mathcal{M}_p : A'B' \to C$ is defined as

$$\mathcal{M}_p(\rho) = \sum_{k=0}^{d-1} p_k (\hat{Z}(k) \otimes \hat{Z}(k)) \rho (\hat{Z}(k) \otimes \hat{Z}(k))^\dagger \tag{54}$$

where $\hat{Z}(k)$ is the generalized Pauli phase operator from (12). We will show that the capacity region for the multiple access phase-flip channel $\mathcal{M}_p$ assisted by entanglement is the collection of all pairs of nonnegative rates (R_1, R_2) which satisfy

$$\begin{aligned} R_1 &\leq 2\log d \\ R_2 &\leq 2\log d \\ R_1 + R_2 &\leq 4\log d - H(p) \end{aligned} \tag{55}$$

Proof First we show that (55) is precisely the region $C_E(\mathcal{M}, \tau, \tau)$, proving achievability. The corresponding θ state is

$$\theta^{ABC} = (I^{AB} \otimes \mathcal{M}_p)(\Phi^{AA'} \otimes \Phi^{BB'}),$$

where $|\Phi\rangle$ is the maximally entangled state (13). It is easy to see that

$$\begin{aligned} H(A) = H(B) = H(\tau) &= \log d \\ H(AC) = H(BC) &= \log d + H(p) \\ H(ABC) &= H(p). \end{aligned} \tag{56}$$

Hence we reach our conclusion

$$\begin{aligned} I(A;C|B) &= 2\log d \\ I(B;C|A) &= 2\log d \\ I(AB;C) &= 4\log d - H(p). \end{aligned} \tag{57}$$

It remains to show that (55) is an upper bound on the capacity region. It is clear from (40) that $R_1 \leq 2H(A)$ and $R_2 \leq 2H(B)$. Hence the first two inequalities in (55). The third makes use of the single-letter upper bound (41) on $R_1 + R_2$. It suffices to show that

$$\max_\rho I(AB;C)_\theta = 4\log d - H(p), \tag{58}$$

where

$$\theta^{ABC} = (I^{AB} \otimes \mathcal{M})(\varphi^{ABA'B'}), \tag{59}$$

and $\varphi^{ABA'B'}$ is a purification of $\rho^{A'B'}$. [1] We need three ingredients. The first is that the maximum in (58) is attained for states $\rho^{A'B'}$ diagonal in the $\{|jl\rangle\}$ basis (see Appendix C for a proof of this fact). Define a Stinespring dilation $U_{\mathcal{M}_p} : A'B' \to CE$ of $\mathcal{M}_p$ as

$$U_{\mathcal{M}_p} = \sum_{jl} |jl\rangle^C |\phi_{jl}\rangle^E \langle jl|^{A'B'} \tag{60}$$

[1] we have already shown that this maximum is achieved for the product state $\rho^{A'B'} = \tau^{A'} \otimes \tau^{B'}$

where

$$|\phi_{jl}\rangle^E = \sum_{k=0}^{d-1} \sqrt{p_k} |k\rangle e^{i2\pi k(j+l)/d}.$$

By the results of Appendix C

$$I(AB;C)_\theta = 2H(\{r_{jl}\}) - H(\sum_{jl} r_{jl}\phi_{jl}), \tag{61}$$

where $\rho = \sum_{jl} r_{jl}|jl\rangle\langle jl|$.

The second ingredient is that $I(AB;C)_\theta$ is a concave function of ρ and hence has a unique local optimum. This is because for *degradable channels* [9] such as $\mathcal{M}_p$, the coherent information $I(AB\rangle C) := I(AB;C) - H(A)$ is a concave function of input density matrix ρ [23]. Since $H(A)$ is also concave we conclude that $I(AB;C)$ is concave.

The third ingredient is to use the method of Lagrange multipliers to find a local optimum for $I(AB;C)_\theta$. We need to optimize

$$f(\{r_{jl}\}) = 2H(\{r_{jl}\}) - H(\sum_{jl} r_{jl}\phi_{jl}) - \lambda \sum_{jl} r_{jl},$$

with Lagrange multiplier λ. Differentiating with respect to the r_{jl} gives d^2 simultaneous equations. By inspection, $r_{jl} = 1/d^2$ is a solution to this system of equations. The second ingredient ensures that this is in fact the global maximum. Thus

$$\max_\rho I(AB;C)_\theta = 2H(\{\frac{1}{d^2}\}) - H(\frac{1}{d^2}\sum_m \phi_m) = 4\log d - H(p),$$

as claimed.

5 A hierarchy of QMAC resource inequalities

In this section we phrase our result using the theory of resource inequalities developed in [7]. The multiple access channel $\mathcal{M} : A'B' \to C$ assisted by some rate E_1 of ebits shared between Alice and Charlie and some rate E_2 of ebits shared between Bob and Charlie, was used to enable a rate R_1 bits of communication between Alice and Charlie and a rate R_2 bits of communication between Bob and Charlie. This is written as

$$\langle \mathcal{M} \rangle + E_1\,[q\,q]_{AC} + E_2\,[q\,q]_{BC} \geq R_1\,[c \to c]_{AC} + R_2\,[c \to c]_{BC}.$$

Without accounting for entanglement consumption (i.e. setting $E_1 = E_2 = \infty$) the above resource inequality holds iff $(R_1, R_2) \in C_E(\mathcal{M})$, with $C_E(\mathcal{M})$ given by Theorem 2. The "if" direction, i.e. the direct coding theorem, followed from the "corner points"

$$\langle \mathcal{M} \rangle + H(A)\,[q\,q]_{AC} + H(B)\,[q\,q]_{BC} \geq I(A;C)\,[c \to c]_{AC} + I(B;CA)\,[c \to c]_{BC} \tag{62}$$

and

$$\langle \mathcal{M} \rangle + H(A)\,[q\,q]_{AC} + H(B)\,[q\,q]_{BC} \geq I(A;CB)\,[c \to c]_{AC} + I(B;C)\,[c \to c]_{BC}. \tag{63}$$

All the entropic quantities are defined relative to the state θ^{ABC} defined in (39).

Just as in the single user case (cf. rule O in [7]), the protocol can be made coherent, replacing $[c \to c]$ by $\frac{1}{2}([q\,q] + [q \to q])$. Cancelling terms on both sides gives "father" protocols for the QMAC

$$\begin{aligned} &\langle \mathcal{M} \rangle + \frac{1}{2}\,I(A;BE)\,[q\,q]_{AC} + \frac{1}{2}\,I(B;E)\,[q\,q]_{BC} \\ &\geq \frac{1}{2}\,I(A;C)\,[q \to q]_{AC} + \frac{1}{2}\,I(B;CA)\,[q \to q]_{BC} \end{aligned} \tag{64}$$

and

$$\begin{aligned}&\langle\mathcal{M}\rangle+\frac{1}{2}\,I(A;E)\,[q\,q]_{AC}+\frac{1}{2}\,I(B;AE)\,[q\,q]_{BC}\\&\geq\frac{1}{2}\,I(A;CB)\,[q\to q]_{AC}+\frac{1}{2}\,I(B;C)\,[q\to q]_{BC},\end{aligned}\tag{65}$$

where the entropic quantities are now defined with respect to a purification θ^{ABCE} of θ^{ABC}.

Applying $[q\to q]\geq[qq]$ to the above equations gives

$$\langle\mathcal{M}\rangle\geq I(A\rangle C)\,[q\to q]_{AC}+\frac{1}{2}\,I(B\rangle CA)\,[q\to q]_{BC}\tag{66}$$

and

$$\langle\mathcal{M}\rangle\geq I(A\rangle BC)\,[q\to q]_{AC}+\frac{1}{2}\,I(B\rangle C)\,[q\to q]_{BC}.\tag{67}$$

These equations are of the form

$$\langle\mathcal{M}\rangle\geq Q_1\,[q\to q]_{AC}+Q_2\,[q\to q]_{BC}.\tag{68}$$

The optimal set of pairs (Q_1,Q_2) satisfying (68) was found in [23], [14]. Equations (66) and (67) recover the "corner points" of the corresponding capacity region.

Coherifying only Bob's resources in equation (62) gives

$$\langle\mathcal{M}\rangle+H(A)\,[q\,q]_{AC}\geq I(A;C)\,[c\to c]_{AC}+I(B\rangle CA)\,[q\to q]_{BC}.$$

Consider $\mathcal{M}$ of a special $\{cq\to q\}$ form in which Alice's input is dephased before being sent though the channel. The arguments from Section 3.2 apply here to show that the Alice-Charlie entanglement is not needed. Thus we recover another coding theorem proven in [23] which characterizes the pairs (R_1,Q_2) for which

$$\langle\mathcal{M}\rangle\geq R_1\,[c\to c]_{AC}+Q_2\,[q\to q]_{BC}.$$

We can also recover the result of Winter [21] which solves

$$\langle\mathcal{M}\rangle\geq R_1\,[c\to c]_{AC}+R_2\,[c\to c]_{BC}.$$

for $\{cc\to q\}$ channels $\mathcal{M}$. We just apply the argument from Section 3.2 to remove the need for any entanglement assistance.

Ultimately we would like to solve

$$\langle\mathcal{M}\rangle\geq Q_1\,[q\to q]_{AC}+E_1\,[q\,q]_{AC}+R_1\,[c\to c]_{AC}+Q_2\,[q\to q]_{BC}+E_2\,[q\,q]_{BC}+R_2\,[c\to c]_{BC},$$

where the 6 rates may be positive or negative. The single user case $Q_2=E_2=R_2=0$ was solved in [8].

Acknowledgements We are grateful to Jon Yard for valuable discussions on generalized dephasing channels. ID and MH were supported by the NSF 05-501 grant no. 0524811. AW was supported by EU grant RESQ (no. IST-2001-37559) and the U.K. Engineering and Physical Sciences Resesrch Council's "QIP IRC".

A Proof of Packing Lemma

We need the following lemma from [11].

Lemma 5 (Hayashi, Nagaoka) *For any operators* $0 \leq S \leq I$ *and* $T \geq 0$*, we have*

$$I - \sqrt{S+T}^{-1} S \sqrt{S+T}^{-1} \leq 2(I-S) + 4T.$$

We are now ready to prove the packing lemma, along lines suggested by the work [11].

Proof Define a collection of pairwise independent random variables $\{M_k : k \in [N]\}$, each distributed according to λ. Set $f(k) = M_k$. Define $M^N = M_1 \dots M_N$. Define $\overline{p}_e = \frac{1}{N}\sum_{k=1}^{N} p_e(k)$ with

$$p_e(k) = \operatorname{Tr} \sigma_{f(k)}(I - \Lambda_k).$$

We shall construct a POVM $\{\Lambda_k\}$ such that

$$\mathbb{E}_{M^N} \overline{p}_e \leq 2(\epsilon + \sqrt{8\epsilon}) + \gamma$$

The POVM elements $\{\Lambda_k\}$ are constructed by the so-called *square root measurement*[12, 17]

$$\Lambda_k = \Big(\sum_l \Upsilon_{f(l)}\Big)^{-\frac{1}{2}} \Upsilon_{f(k)} \Big(\sum_l \Upsilon_{f(l)}\Big)^{-\frac{1}{2}}$$

with

$$\Upsilon_m = \Pi \Pi_m \Pi.$$

Invoking lemma 5, we can now place an upper bound on the probability error:

$$\overline{p}_e \leq \frac{1}{N} \sum_{k=1}^{N} [2(1 - \operatorname{Tr} \sigma_{f(k)} \Upsilon_{f(k)}) + 4 \sum_{l \neq k} \operatorname{Tr} \sigma_{f(k)} \Upsilon_{f(l)}]. \tag{69}$$

Lemma **??** and property (18) give

$$\|\Pi \sigma_m \Pi - \sigma_m\|_1 \leq \sqrt{8\epsilon}. \tag{70}$$

By property (17) and (70)

$$\begin{aligned} \operatorname{Tr} \sigma_m \Upsilon_m &\geq \operatorname{Tr} \sigma_m \Pi_m - \|\Pi \sigma_m \Pi - \sigma_m\|_1 \\ &\geq 1 - \epsilon - \sqrt{8\epsilon}. \end{aligned} \tag{71}$$

For $k \neq l$, the random variables M_k and M_l are independent. Thus

$$\begin{aligned} \mathbb{E}_{M^N}(\operatorname{Tr} \sigma_{f(k)} \Upsilon_{f(l)}) &= \operatorname{Tr}(\Pi \mathbb{E} \sigma_{M_k} \Pi \ \mathbb{E} \Pi_{M_l}) \\ &\leq D^{-1} \mathbb{E} \operatorname{Tr} \Pi \Pi_{M_l} \\ &\leq d/D. \end{aligned} \tag{72}$$

The first inequality follows from $\mathbb{E}\, \sigma_{M_k} = \sigma$ and property (19). The second follows from $\Pi \leq I$ and property (20). Taking the expectation of (69), and incorporating (71) and (72) gives

$$\begin{aligned} \mathbb{E}_{M^N} \overline{p}_e &\leq 2(\epsilon + \sqrt{8\epsilon}) + 4(N-1)d/D, \\ &\leq 2(\epsilon + \sqrt{8\epsilon}) + 4Nd/D \\ &= 2(\epsilon + \sqrt{8\epsilon}) + 4\gamma =: \epsilon'. \end{aligned} \tag{73}$$

Two more standard steps are needed.

i) Derandomization. There exists at least one particular value m^N of the string M^N such that for which $\overline{p}_e$ is at least as small as the expectation value. Thus

$$\overline{p}_e(\mathcal{C}) \leq \epsilon'. \tag{74}$$

ii) Average to maximal error probability. Since

$$\overline{p}_e = \frac{1}{N} \sum_{k \in N} p_e(k) \leq \epsilon',$$

then $p_e(k) \leq 2\epsilon'$ for at least half the indices k. Throw the others away and redefine f, N and γ accordingly. This further changes the error estimate to

$$4(\epsilon + \sqrt{8\epsilon}) + 16\gamma.$$

■

B Proofs of properties (30)-(33)

I. Proof of property (30).
Define $\check{P}$ to be the complement of the projector P. That is $\check{P} = I - P$.

$$\begin{aligned} \Pi &= \Pi^n_{\mathcal{N}(\rho),\delta} \otimes \Pi^n_{\rho,\delta} \\ &= (I - \check{\Pi}^n_{\mathcal{N}(\rho),\delta}) \otimes (I - \check{\Pi}^n_{\rho,\delta}) \\ &= I \otimes I - I \otimes \check{\Pi}^n_{\rho,\delta} - \check{\Pi}^n_{\mathcal{N}(\rho),\delta} \otimes I + \check{\Pi}^n_{\mathcal{N}(\rho),\delta} \otimes \check{\Pi}^n_{\rho,\delta} \\ &\geq I \otimes I - I \otimes \check{\Pi}^n_{\rho,\delta} - \check{\Pi}^n_{\mathcal{N}(\rho),\delta} \otimes I. \end{aligned} \tag{75}$$

Therefore

$$\begin{aligned} \operatorname{Tr} \sigma^{BB'}_{s^a} \Pi &\geq \operatorname{Tr} \sigma_{s^a} - \operatorname{Tr} \sigma_{s^a} (I \otimes \check{\Pi}^n_{\rho,\delta}) - \operatorname{Tr} \sigma_{s^a} (\check{\Pi}^n_{\mathcal{N}(\rho),\delta} \otimes I) \\ &= 1 - \operatorname{Tr} [\sigma^{B'}_{s^a} \check{\Pi}^n_{\rho,\delta}] - \operatorname{Tr} [\sigma^{B}_{s^a} \check{\Pi}^n_{\mathcal{N}(\rho),\delta}] \\ &\geq 1 - 2\epsilon, \end{aligned} \tag{76}$$

the last line by a double application of (8).

II. Proof of property (31).
By (25) and (29),

$$\begin{aligned} \operatorname{Tr} \sigma_{s^a} \Pi_{s^a} &= \operatorname{Tr} \theta^{\otimes n} \Pi^n_{\theta,\delta} \\ &\geq 1 - \epsilon, \end{aligned} \tag{77}$$

The last line follows from (8).

III. Proof of property (32).

$$\operatorname{Tr} \Pi_{s^a} = \operatorname{Tr} \Pi^n_{\theta,\delta} \leq 2^{n[H(AB)_\theta + c\delta]}. \tag{78}$$

The inequality follows from (10).

IV. Proof of property (33).

Because of (11), we can bound the density operator τ_α^n by

$$\tau_\alpha^n = \frac{\Pi^n_{t(\alpha)}}{\operatorname{Tr}\Pi^n_{t(\alpha)}} \leq 2^{-n[H(\rho)-\eta(\delta)]}\Pi^n_{\rho,\delta}. \tag{79}$$

Then

$$\begin{aligned}
\Pi\sigma\Pi &= (\Pi^n_{\mathcal{N}(\rho),\delta} \otimes \Pi^n_{\rho,\delta}) \left[\sum_\alpha p_\alpha (\mathcal{N}^{\otimes n}(\tau^n_\alpha) \otimes \tau^n_\alpha)\right] (\Pi^n_{\mathcal{N}(\rho),\delta} \otimes \Pi^n_{\rho,\delta}) \\
&= \sum_\alpha p_\alpha \left[(\Pi^n_{\mathcal{N}(\rho),\delta}\mathcal{N}^{\otimes n}(\tau^n_\alpha)\Pi^n_{\mathcal{N}(\rho),\delta}) \otimes (\Pi^n_{\rho,\delta}\tau^n_\alpha\Pi^n_{\rho,\delta})\right] \\
&\leq \left(\Pi^n_{\mathcal{N}(\rho),\delta}\mathcal{N}^{\otimes n}(\sum_\alpha p_\alpha \tau^n_\alpha)\Pi^n_{\mathcal{N}(\rho),\delta}\right) \otimes (2^{-n[H(\rho)-\eta(\delta)]}\Pi^n_{\rho,\delta}) \\
&\leq \left(2^{-n[H(\mathcal{N}(\rho))-c\delta]}\Pi^n_{\mathcal{N}(\rho),\delta}\right) \otimes \left(2^{-n[H(\rho)-\eta(\delta)]}\Pi^n_{\rho,\delta}\right) \\
&= 2^{-n[H(\rho)+H(\mathcal{N}(\rho))-c\delta-\eta(\delta)]}\,\Pi \\
&= 2^{-n[H(A)_\theta+H(B)_\theta-c\delta-\eta(\delta)]}\,\Pi,
\end{aligned} \tag{80}$$

where the first inequality follows from (79) and the second from (9).

C Generalized dephasing channels

We follow the techniques of [23, 22, 9]. Let A' and B be quantum systems of dimension d with respective bases $\{|i\rangle^{A'}\}$ and $\{|i\rangle^B\}$.

The a channel $\mathcal{N} : A' \to B$ is called a generalized dephasing channel if

$$\mathcal{N}(|i\rangle\langle i|^{A'}) = |i\rangle\langle i|^B.$$

We can write down a Stinespring dilation $U_\mathcal{N} : A' \to BE$ for $\mathcal{N}$:

$$U_\mathcal{N} = \sum_i |i\rangle^B|\phi_i\rangle^E\langle i|^{A'},$$

where the $\{|\phi_i\rangle^E\}$ are not necessarily orthogonal. Given $U_\mathcal{N}$, the complementary channel $\mathcal{N}^c$: $A' \to E = \operatorname{Tr}_B \circ U_\mathcal{N}$ acts on some input state $\rho^{A'}$ as

$$\begin{aligned}
\mathcal{N}^c(\rho) &= \operatorname{Tr}_B U_\mathcal{N}(\rho) \\
&= \sum_i \langle i|^B \Big(\sum_{i''i'} |i''\rangle^B|\phi_{i''}\rangle^E\langle i''|^{A'}\rho|i'\rangle^{A'}\langle i'|^B\langle\phi_{i'}|^E\Big)|i\rangle^B \\
&= \sum_i \langle i|\rho|i\rangle\phi_i^E \\
&=: \sum_i r_i\phi_i^E.
\end{aligned} \tag{81}$$

It depends only on the diagonal elements $\{r_i\}$ of ρ expressed in the dephasing basis. When the $\{|\phi_i\rangle^E\}$ are also orthogonal, the channel $\mathcal{N}$ is called *completely dephasing* and is denoted by $\triangle$. It corresponds to performing a projective measurement in the dephasing basis and ignoring the result. The following properties hold:

$$\begin{aligned}
\mathcal{N}^c &= \mathcal{N}^c \circ \triangle \\
\mathcal{N} \circ \triangle &= \triangle \circ \mathcal{N} \\
H(\triangle(\rho)) &\geq H(\rho)
\end{aligned} \tag{82}$$

Define $\theta^{AB} = (I^A \otimes \mathcal{N})\phi^{AA'}$, where $\phi^{AA'}$ is a purification of the input state $\rho^{A'}$.

Lemma 6 *Given a dephasing channel $\mathcal{N} : A' \to B$, the mutual information $I(A;B)_\theta$ is maximal when the input state $\rho^{A'}$ is diagonal in the dephasing basis.*

Proof Since

$$
\begin{aligned}
I(A;B) &= H(A) + H(B) - H(BA) \\
&= H(A) + H(B) - H(E) \\
&= H(\rho) + H(\mathcal{N}(\rho)) - H(\mathcal{N}^c(\rho)) \\
&\leq H(\triangle(\rho)) + H((\triangle \circ \mathcal{N}(\rho)) - H(\mathcal{N}^c \circ \triangle(\rho)) \\
&= H(\triangle(\rho)) + H(\mathcal{N} \circ \triangle(\rho)) - H(\mathcal{N}^c \circ \triangle(\rho))
\end{aligned} \tag{83}
$$

The inequality is saturated when $\rho = \triangle(\rho) = \sum r_i |i\rangle\langle i|$, in which case

$$I(A;B) = 2H(\{r_i\}) - H(\sum_i r_i \phi_i).$$

References

[1] R. Ahlswede. Multi-way communication channels. *Second Intern. Sympos. on Inf. Theory, Thakadsor*, 1971.

[2] C. H. Bennett, P. W. Shor, J. A. Smolin, and A. Thapliyal. Entanglement-assisted capacity of a quantum channel and the reverse shannon theorem. *IEEE Trans. Inf. Theory*, 48(10):2637–2655, 2002. quant-ph/0106052.

[3] C. H. Bennett and S. J. Wiesner. Communication via one- and two-particle operators on einstein-podolsky-rosen states. *Phys. Rev. Lett.*, 69:2881–2884, 1992.

[4] T. M. Cover and J. A. Thomas. *Elements of Information Theory.* Series in Telecommunication. John Wiley and Sons, New York, 1991.

[5] I. Csiszár and J. Körner. *Information Theory: coding theorems for discrete memoryless systems.* Academic Press, New York–San Francisco–London, 1981.

[6] E. B. Davies and J. T. Lewis. An operational approach to quantum probability. *Comm. Math. Phys.*, 17:239–260, 1970.

[7] I. Devetak, A. W. Harrow, and A. J. Winter. A family of quantum protocols. *Phys. Rev. Lett.*, 93:230504, 2004. quant-ph/0308044.

[8] I. Devetak, P. Hayden, D. W. Leung, and P. Shor. Triple trade-offs in quantum shannon theory, 2005. in preparation.

[9] I. Devetak and P. W. Shor. The capacity of a quantum channel for simultaneous transmission of classical and quantum information, 2003. quant-ph/0311131.

[10] A. W. Harrow and H. K. Lo. A tight lower bound on the classical communication cost of entanglement dilution. *IEEE Trans. Inf. Theory*, 50(2):319–327, 2004. quant-ph/0204096.

[11] M. Hayashi and H. Nagaoka. General formulas for capacity of classical-quantum channels. *IEEE Trans. Inf. Theory*, 49(7):1753–1768, 2003. quant-ph/0206186.

[12] A. S. Holevo. The capacity of the quantum channel with general signal states. *IEEE Trans. Inf. Theory*, 44(1):269–273, 1998. quant-ph/9611023.

[13] A. S. Holevo. On entanglement assisted classical capacity. *J. Math. Phys.*, 43(9):4326–4333, 2002. quant-ph/0106075.

[14] M. Horodecki, J. Oppenheim, and A. Winter. Partial quantum information. *Nature*, 436:673–676, 2005. quant-ph/0505062.

[15] H. Liao. *Multiple access channels*. PhD thesis, University of Hawaii, 1972.

[16] E. H. Lieb and M. B. Ruskai. Proof of strong subadditivity of quantum-mechanical entropy. *J. Math. Phys.*, 14:1938–1941, 1973.

[17] B. Schumacher and M. D. Westmoreland. Sending classical information via noisy quantum channels. *Phys. Rev. A*, 56(1):131–138, 1997. quant-ph/9604023.

[18] C. E. Shannon. A mathematical theory of communication. *Bell System Tech. Jnl.*, 27:379–423, 623–656, 1948.

[19] P. W. Shor. Fault tolerant quantum computation. In *Proceedings of the 37th Symposium on the Foundations of Computer Science*, pages 56–65, Los Alamitos, CA, 1996. IEEE. quant-ph/9605011.

[20] A. Winter. Coding theorem and strong converse for quantum channels. *IEEE Trans. Inf. Theory*, 45(7):2481–2485, 1999.

[21] A. Winter. The capacity of the quantum multiple access channel. *IEEE Trans. Inf. Theory*, 47:3059–3065, 2001. quant-ph/9807019.

[22] J. Yard. *Simultaneous classical-quantum capacities of quantum multiple access channels*. PhD thesis, Stanford University, 2005. quant-ph/0506050.

[23] J. Yard, I. Devetak, and P. Hayden. Capacity theorems for quantum multiple access channels: Classical-quantum and quantum-quantum capacity regions, 2005. quant-ph/0501045.